

SEQUENCE LISTING

<110> Miao, Carol
Kay, Mark

<120> Liver-Specific Gene Expression Cassettes, and Methods of Use

<130> UOFW-1-17396

<150> US 60/212,902

<151> 2000-06-20

<160> 18

<170> PatentIn version 3.0

<210> 1

<211> 1438

<212> DNA

<213> HomoSapien

<400> 1

g t t t t g t t t c c t t t t t a a a a t a c a t t g a g t a t g c t t g c c t t t a g a t a t a g a a t a t c t g 60
a t g c t g t c t t c t t c a c t a a a t t t t g a t t a c a t g a t t t g a c a g c a a t a t t g a a g a g t c t a a 120
c a g c c a g c a c g c a g g t t g g t a a g t a c t g g t t c t t t g t t a g c t a g g t t t t c t t c t t c t t c a 180
t t t t t a a a a c t a a t a g a t c g a c a a t g c t t a t g a t g c a t t t a t g t t t a a t a a a c a c t g t t 240
c a g t t c a t g a t t t g g t c a t g t a a t t c o t g t t a g a a a a c a t t c a t c t c e c t t g g t t t a a a a a 300
a a t t a a a a g t g g g a a a c a a a g a a t a g c a g a a t a t a g t g a a a a a a a t a a c c a c a t t a t 360
t t t t g t t t g g a c t t a c c a c t t t g a a a t c a a a a t g g g a a a c a a a a g c a c a a a c a a t g g c c t 420
t a t t t a c a c a a a a g t c t g a t t t a a g a t a t a t g a c a t t t c a a g g t t t c a g a a g t a t g t a 480
a t g a g g t g t g t c t c t a a t t t t t a a a t t a t a t a t c t t c a a t t t a a a g t t t t a g t t a a a a c 540
a t a a a g a t t a a c t t t t c a t t a g c a a g c t g t t a g t t a t c a c c a a a g c t t t t c a t g g a t t a g 600
g a a a a a t c a t t t t g t c t c t a t g t c a a a c a t c t t g g a g t t g a t a t t t g g g g a a a c a c a a t 660
a c t c a g t t g a g t t c c c t a g g g g a g a a a a g c a a g c t t a a g a a t t g a c a t a a a g a g t a g g a a 720
g t t a g c t a a t g c a a c a t a t a t c a c t t t g t t t t t c a c a a c t a c a g t g a c t t t a t g t a t t t 780
c c c a g a g g a a g g c a t a c a g g g a a g a a a t t a t c c c a t t t g g a c a a a c a g c a t g t t t c t c a c a 840
g g a a g c a t t t a t c a c a c t t a c t t g t c a a c t t t c t a g a a t c a a a t c t a g t a g c t g a c a g t a 900

"COLE" PROTEIN

Glu	Ala	Arg	Glu	Val	Phe	Glu	Asn	Thr	Glu	Arg	Thr	Thr	Glu	Phe	Trp		
		75					80					85					
aag	cag	tat	ggt	gat	gga	gat	cag	tgt	gag	tcc	aat	cca	tgt	tta	aat	341	
Lys	Gln	Tyr	Val	Asp	Gly	Asp	Gln	Cys	Glu	Ser	Asn	Pro	Cys	Leu	Asn		
	90					95					100						
ggc	ggc	agt	tgc	aag	gat	gac	att	aat	tcc	tat	gaa	tgt	tgg	tgt	ccc	389	
Gly	Gly	Ser	Cys	Lys	Asp	Asp	Ile	Asn	Ser	Tyr	Glu	Cys	Trp	Cys	Pro		
105					110					115					120		
ttt	gga	ttt	gaa	gga	aag	aac	tgt	gaa	tta	gat	gta	aca	tgt	aac	att	437	
Phe	Gly	Phe	Glu	Gly	Lys	Asn	Cys	Glu	Leu	Asp	Val	Thr	Cys	Asn	Ile		
			125						130					135			
aag	aat	ggc	aga	tgc	gag	cag	ttt	tgt	aaa	aat	agt	gct	gat	aac	aag	485	
Lys	Asn	Gly	Arg	Cys	Glu	Gln	Phe	Cys	Lys	Asn	Ser	Ala	Asp	Asn	Lys		
			140					145						150			
gtg	ggt	tgc	tcc	tgt	act	gag	gga	tat	cga	ctt	gca	gaa	aac	cag	aag	533	
Val	Val	Cys	Ser	Cys	Thr	Glu	Gly	Tyr	Arg	Leu	Ala	Glu	Asn	Gln	Lys		
		155					160						165				
tcc	tgt	gaa	cca	gca	gtg	cca	ttt	cca	tgt	gga	aga	gtt	tct	ggt	tca	581	
Ser	Cys	Glu	Pro	Ala	Val	Pro	Phe	Pro	Cys	Gly	Arg	Val	Ser	Val	Ser		
	170					175					180						
caa	act	tct	aag	ctc	acc	cgt	gct	gag	gct	ggt	ttt	cct	gat	gtg	gac	629	
Gln	Thr	Ser	Lys	Leu	Thr	Arg	Ala	Glu	Ala	Val	Phe	Pro	Asp	Val	Asp		
185				190						195					200		
tat	gta	aat	tct	act	gaa	gct	gaa	acc	att	ttg	gat	aac	atc	act	caa	677	
Tyr	Val	Asn	Ser	Thr	Glu	Ala	Glu	Thr	Ile	Leu	Asp	Asn	Ile	Thr	Gln		
			205						210					215			
agc	acc	caa	tca	ttt	aat	gac	ttc	act	cgg	ggt	ggt	ggt	gga	gaa	gat	725	
Ser	Thr	Gln	Ser	Phe	Asn	Asp	Phe	Thr	Arg	Val	Val	Gly	Gly	Glu	Asp		
			220					225					230				
gcc	aaa	cca	ggt	caa	ttc	cct	tgg	cag	ggt	ggt	ttg	aat	ggt	aaa	ggt	773	
Ala	Lys	Pro	Gly	Gln	Phe	Pro	Trp	Gln	Val	Val	Leu	Asn	Gly	Lys	Val		
		235					240						245				
gat	gca	ttc	tgt	gga	ggc	tct	atc	ggt	aat	gaa	aaa	tgg	att	gta	act	821	
Asp	Ala	Phe	Cys	Gly	Gly	Ser	Ile	Val	Asn	Glu	Lys	Trp	Ile	Val	Thr		
	250					255					260						
gct	gcc	cac	tgt	ggt	gaa	act	ggt	ggt	aaa	att	aca	ggt	gtc	gca	ggt	869	
Ala	Ala	His	Cys	Val	Glu	Thr	Gly	Val	Lys	Ile	Thr	Val	Val	Ala	Gly		
265					270					275					280		

PROTEIN "TUBULIN"

gaa cat aat att gag gag aca gaa cat aca gag caa aag cga aat gtg 917
Glu His Asn Ile Glu Glu Thr Glu His Thr Glu Gln Lys Arg Asn Val
285 290 295

att cga att att cct cac cac aac tac aat gca gct att aat aag tac 965
Ile Arg Ile Ile Pro His His Asn Tyr Asn Ala Ala Ile Asn Lys Tyr
300 305 310

aac cat gac att gcc ctt ctg gaa ctg gac gaa ccc tta gtg cta aac 1013
Asn His Asp Ile Ala Leu Leu Glu Leu Asp Glu Pro Leu Val Leu Asn
315 320 325

agc tac gtt aca cct att tgc att gct gac aag gaa tac acg aac atc 1061
Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp Lys Glu Tyr Thr Asn Ile
330 335 340

ttc ctc aaa ttt gga tct ggc tat gta agt ggc tgg gga aga gtc ttc 1109
Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser Gly Trp Gly Arg Val Phe
345 350 355 360

cac aaa ggg aga tca gct tta gtt ctt cag tac ctt aga gtt cca ctt 1157
His Lys Gly Arg Ser Ala Leu Val Leu Gln Tyr Leu Arg Val Pro Leu
365 370 375

ggt gac cga gcc aca tgt ctt cga tct aca aag ttc acc atc tat aac 1205
Val Asp Arg Ala Thr Cys Leu Arg Ser Thr Lys Phe Thr Ile Tyr Asn
380 385 390

aac atg ttc tgt gct ggc ttc cat gaa gga ggt aga gat tca tgt caa 1253
Asn Met Phe Cys Ala Gly Phe His Glu Gly Gly Arg Asp Ser Cys Gln
395 400 405

gga gat agt ggg gga ccc cat gtt act gaa gtg gaa ggg acc agt ttc 1301
Gly Asp Ser Gly Gly Pro His Val Thr Glu Val Glu Gly Thr Ser Phe
410 415 420

tta act gga att att agc tgg ggt gaa gag tgt gca atg aaa ggc aaa 1349
Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu Cys Ala Met Lys Gly Lys
425 430 435 440

tat gga ata tat acc aag gta tcc cgg tat gtc aac tgg att aag gaa 1397
Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr Val Asn Trp Ile Lys Glu
445 450 455

aaa aca aag ctc act t 1413
Lys Thr Lys Leu Thr
460

<210> 3
<211> 461

<212> PRT

<213> HomoSapien

<400> 3

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn
35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys
50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn
65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly Asp Gln
85 90 95

Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp Asp Ile
100 105 110

Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys Asn Cys
115 120 125

Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe
130 135 140

Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly
145 150 155 160

Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val Pro Phe
165 170 175

Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala
180 185 190

FOR "FOR" FOR

385

390

395

400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val
405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly
420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser
435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
450 455 460

CGGGGCG
CCGCGCG
CGGGGCG
<210> 4
<211> 771
<212> DNA
<213> HomoSapien

<400> 4
caggctcaga ggcacacagg agtttctggg ctccacctgc ccccttccaa cccctcagtt 60
cccacacctc agcagctggt tgtgtgctgc ctctgaagtc cacactgaac aaacttcagc 120
ctactcatgt ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct 180
ccctgcctgc tgaccttggg gctggggcag aggtcagaga cctctctggg cccatgccac 240
ctccaacatc cactcgaccc cttggaattt cgggtggagag gagcagaggt tgtcctggcg 300
tggtttaggt agtgtgagag ggtccggggt caaaaccact tgctgggtgg ggagtcgtca 360
gtaagtggct atgccccgac cccgaagcct gtttcccat ctgtacaatg gaaatgataa 420
agacgcccat ctgatagggt ttttgtggca aataaacatt tggttttttt gttttgtttt 480
gttttgtttt ttgagatgga ggtttgctct gtcgcccagg ctggagtgca gtgacacaat 540
ctcatctcac cacaacctc cctgcctca gcctccaag tagctgggat tacaagcatg 600
tgccaccaca cctggctaata tttctatttt tagtagagac gggtttctcc atggttggtca 660
gcctcagcct cccaagtaac tgggattaca ggctgtgcc accacacccg gctaattttt 720
tctatttttg acagggacgg ggtttcacca tgttggtcag gtcctctag a 771

<210> 5
<211> 418
<212> DNA
<213> HomoSapien

<400> 5
ggatcttgct accagtgga cagccactaa ggattctgca gtgagagcag agggccagct 60
aagtgggtact ctcccagaga ctgtctgact cacgccaccc cctccacctt ggacacagga 120
cgctgtgggtt tctgagccag gtacaatgac tcctttcggt aagtgcagtg gaagctgtac 180
actgcccagg caaagcgtcc gggcagcgtg ggcgggagac tcagatccca gccagtggac 240
ttagcccttg tttgctctc cgataactgg ggtgacctg gttaatattc accagcagcc 300
tccccggtg cccctctgga tccactgctt aaatacggac gaggacaggg ccctgtctcc 360
tcagcttcag gcaccaccac tgacctgga cagtgaatga tccccctgat ctgcggcc 418

<210> 6
<211> 282
<212> DNA
<213> Bos taurus

<400> 6
cccgggggat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgccctcc 60
cccgtgcctt ccttgacctt ggaaggtgcc actcccactg tcctttccta ataaaatgag 120
gaaattgcat cgcattgtct gagtaggtgt cattctattc tggggggtgg ggtggggcag 180
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 240
atggcttctg aggcggaaag aaccagctgg ggctcgagat cc 282

<210> 7
<211> 1707
<212> DNA
<213> HomoSapien

<400> 7
aatgaaagat ggatttcaa ggtaattca ttggaattga aaattaacag ggcctctcac 60
taactaatca ctttcccatc ttttggtaga tttgaatata tacattctat gatcattgct 120
ttttctcttt acaggggaga atttcatatt ttacctgagc aaattgatta gaaaatggaa 180

CCDS000001.1

<210> 8
<211> 154
<212> DNA
<213> HomoSapien

<400> 8
gtttgtgtgc tgcctctgaa gtccacactg aacaaacttc agcctactca tgtccctaaa 60
atgggcaaac attgcaagca gcaaacagca aacacacagc cctccctgcc tgetgacctt 120
ggagctgggg cagaggtcag agacctctct gggc 154

<210> 9
<211> 328
<212> DNA
<213> Homo Sapien

<400> 9
caggctcaga ggcacacagg agtttctggg ctcaccctgc ccccttccaa cccctcagtt 60
cccctcctcc agcagctggt tgtgtgctgc ctctgaagtc cacactgaac aaacttcagc 120
ctactcatgt ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct 180
ccctgcctgc tgaccttggg gctggggcag aggtcagaga cctctctggg cccatgccac 240
ctccaacatc cactcgacct cttggaattt cgggtggagag gagcagaggt tgtcctggcg 300
tggtttaggt agtgtgagag ggtccggg 328

<210> 10
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> HNF-1 Hepatic Nuclear Factor Binding Site Consensus Sequence

<400> 10
tgtaacag 8

<210> 11
<211> 22
<212> DNA

CCDS10000.1

<213> Artificial Sequence

<220>

<221> misc_feature

<223> HNF-1 Alternative Hepatic Nuclear Factor Binding Site Consensus Sequence

<400> 11

cacggataaa tatgaacctt gg

22

<210> 12

<211> 11

<212> DNA

<213> Artificial Sequence

<220>

<221> misc_feature

<223> HNF-3alpha Hepatic Nuclear Binding Site Consensus Sequence

<400> 12

tattgayttw g

11

<210> 13

<211> 10

<212> DNA

<213> Artificial Sequence

<220>

<221> misc_feature

<223> HNF-3beta Hepatic Nuclear Binding Site Consensus Sequence

<400> 13

atattgattt

10

<210> 14

<211> 10

<212> DNA

<213> Artificial Sequence

<220>

<221> misc_feature

<223> HNF-4 Hepatic Nuclear Binding Site Consensus Sequence

<400> 14

FORF90" FCF9860

aagycaayha 10

- <210> 15
- <211> 10
- <212> DNA
- <213> Artificial Sequence

- <220>
- <221> misc_feature
- <223> HNF-6 Hepatic Nuclear Binding Site Consensus Sequence

<400> 15
aatcaattt 10

- <210> 16
- <211> 13
- <212> DNA
- <213> Artificial Sequence

- <220>
- <221> misc_feature
- <223> HNF-6 Alternative Hepatic Nuclear Binding Site Consensus Sequence

<400> 16
attattgata aaa 13

- <210> 17
- <211> 30
- <212> DNA
- <213> Artificial Sequence

- <220>
- <221> misc_feature
- <223> Human Factor IX cDNA Primer Consensus Sequence

<400> 17
gatggagatc agtgtgagtc caatccatgt 30

- <210> 18
- <211> 30
- <212> DNA
- <213> Artificial Sequence

- <220>

FOR 90 " FOR 90

<221> misc_feature

<223> Human Factor IX cDNA Primer Consensus Sequence

<400> 18

agccacttac atagccagat ccaaatttga

30

09884901.051301